Large-scale 21-cm Cosmology with LOFAR and AARTFAAC

Gehlot, Bharat Kumar

1 Introduction

1.1 An overview

Understanding the formation and evolution of structures during different phases in the history of the Universe is one of the major frontiers in present-day astronomy and cosmology. Around 400,000 years after the Big Bang, matter decoupled from radiation, leaving a relic signature called the Cosmic Microwave Background (CMB) radiation. Observations of the CMB radiation by the WMAP (Wilkinson Microwave Anisotropy Probe) and Planck missions have shed light on the formation of the Universe in the Big-Bang cosmology paradigm (Hinshaw et al. 2013; Planck Collaboration et al. 2016b). These observations provide a plethora of information about the major constituents of the Universe: baryonic (visible) matter, dark matter, and dark energy, as well as provide constraints on the structure formation in the infant Universe. After matter decoupled from radiation, tiny matter density fluctuations grew under gravitational instability to form the first structures of dark and baryonic matter, which are thought to have evolved into a cosmic web of sheets, filaments, and clusters as we observe them today (Springel et al. 2005). On the other hand, observations of galaxies and clusters of galaxies in the local universe also provide a wealth of information, improving our understanding of the astrophysics of galaxies and structure formation in the local Universe (Cole et al. 2005). A combination of these observations generates a paradigm of structure formation in the Universe. Complex objects such as stars, galaxies, and clusters of galaxies that surround us today, are formed inside the overdense regions that grew out of tiny fluctuations in the otherwise very smooth matter density that emerged from the Big Bang.

However, a substantial hiatus still exists in our current knowledge of the Universe. In particular, the cosmic *Dark Ages* which started after matter decoupled from radiation, the Cosmic Dawn (CD) when the first luminous objects appeared and heated the Inter-Galactic Medium (IGM), and the Epoch of Reionization (EoR) during which ionizing radiation from the first generation of stars and galaxies started to ionize the IGM, are all poorly explored. The CD and EoR are pivotal periods in the history of the Universe as they bear the imprint of structure formation and evolution in the early stages of evolution of the Universe. Apart from the CMB, so far no direct detection of the emission of the IGM exist for these intervening epochs when the first stars and galaxies are thought to have formed. Our current knowledge of star and galaxy formation is based on nearby objects which evolve in a completely different environment, compared to theses first-generation luminous objects. Hence it is crucial to study the properties of these first stars and galaxies during the CD and EoR. Several experiments aim to study these cosmic epochs via the 21-cm line of neutral hydrogen, using, or planning to use, a wide range of current

and next-generation radio telescopes. This thesis presents various techniques to study and mitigate the challenges in observing these periods and explores the possibility of observing the Cosmic Dawn with the LOFAR and AARTFAAC radio telescopes, and present results obtained from these experiments.

In this chapter, we first go through different transitional periods in the IGM history of the Universe and place the CD and EoR on this cosmic timeline (section 1.2). We briefly summarise the current observational constraints on the CD and EoR in section 1.3. In section 1.4, we present the redshifted 21-cm line of neutral hydrogen as a new observational probe of the CD and EoR and provide a brief overview of experiments aiming to detect its signal. We provide a brief description of the LOFAR-EoR key science project and the AARTFAAC Cosmic Explorer, and a brief technical overview of the radio telescopes used by these projects, in sections 1.4.2 and 1.4.3. Section 1.5 presents an overview of the key challenges faced by these 21-cm experiments. Finally, this introductory chapter concludes with a brief overview and the major objectives of the chapters in this thesis (section 1.6).

1.2 A Short Chronology of the Universe

The Big Bang (BB) model of the origin and evolution of the Universe is broadly accepted. According to this model, the Universe started 13.8 billion years ago, likely from a quantum-gravity singularity. Roughly 10^{-43} seconds after the BB, the Universe was homogeneous to an extremely high level, isotropic and in a hot, dense and quantum state. The Universe is thought to have gone through a rapid expansion at an exponential rate via a process called inflation, and its size grew by a factor of about 10^{26}. This expansion also inflated inhomogeneities in the matter density, caused by quantum fluctuations (Kolb & Turner 1990), by a large factor. Soon after this, matter annihilated with antimatter, and because of a minute asymmetry between the two, left an excess of baryons. After this, the Universe continued to expand and cool down. One second after the BB, the Universe was filled with a hot soup of neutrons, protons, electrons, positrons, photons, and neutrinos with a temperature of around 10^{10} K. Three minutes after the BB, the Universe cooled down sufficiently for neutrons and protons to combine via fusion and form light nuclei such as deuterium, helium and their isotopes. This process of the formation of light elements, via fusion, is called primordial nucleosynthesis.

The universe at this stage was still hot, dense and full of subatomic particles and light nuclei. The Universe was opaque and acted similarly to a blackbody due to the very high scattering cross-section between photons and electrons. As the universe expanded further, it cooled down, decreasing its density. This decrease, in turn, reduced scattering. Finally, after $\sim 380,000$ years, the Uni-

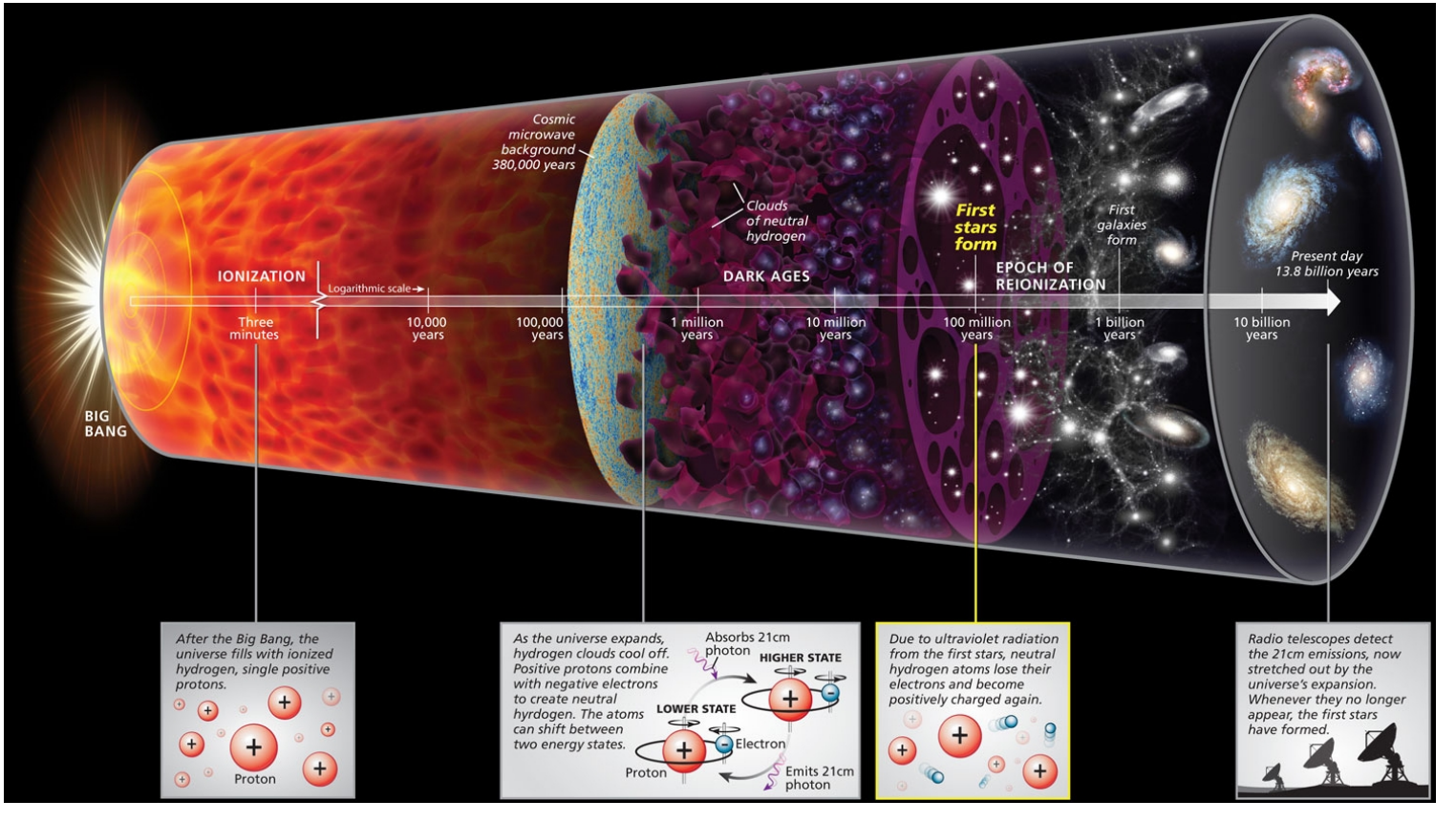

Figure 1.1 – Cartoon depicting a timeline of the Universe as per current understanding of the structure formation in the Universe. Image credit: Roen Kelly/*Discover magazine*.

verse cooled below 3000 K, which allowed ions and electrons to recombine and form neutral hydrogen and helium via the process of recombination. At this point ($z \sim 1100$), photons decoupled from baryons and the Universe became largely transparent. The relic radiation from this last scattering surface has a blackbody spectrum and is known as the Cosmic Microwave Background (CMB) radiation. Today, the peak of the blackbody spectrum is redshifted and into the microwave part of the electromagnetic spectrum, observed e.g., by the WMAP and Planck space missions.

Recombination ushered the universe into the cosmic *Dark Ages* ($1100 \gtrsim z \gtrsim 30$), during which the Universe was mostly dark, neutral and devoid of any luminous objects other than the fading CMB radiation. During this period, small perturbations in the matter density left over from inflation, grew further under the influence of gravitational instability to form first non-linear structures of dark and baryonic matter (dark matter 'halos' and stars and galaxies inside of them). The neutral gas falling into these halos is thought to have formed the first generation of stars and have accreted onto black holes, which together later formed the first galaxies. Formation of the first stars around 400 Myr after the BB, ended the *dark ages* and marked the beginning of the Cosmic Dawn (CD; $30 \gtrsim z \gtrsim 12$). The ultraviolet and X-ray radiation from these stars and black holes heated the cold neutral hydrogen in the surrounding IGM, and during the transitional Epoch of Reionization (EoR; $12 \gtrsim z \gtrsim 6$), radiation ionised most of the IGM. The EoR lasted for about 600 Myr until the hydrogen in the IGM phase transitioned nearly fully from being neutral to ionized. For a detailed theoretical overview of the astrophysics during these epochs, the reader is referred to Furlanetto, Oh & Briggs (2006); Zaroubi (2013)

1.3 Current status of affairs

Although the narrative presented above provides a fiducial model of the structure formation in the Universe in the first 1 Gyr after the BB, several key questions remain unanswered: e.g. *What was the timing and duration of CD and EoR? Exactly when did these first stars form and what were their properties (e.g., mass, luminosity function)? What sources were responsible for heating and ionizing the IGM? What were the properties of the first galaxies, e.g. their type, size, mass, luminosity, morphology, etc?*

Direct observations of the IGM in these epochs are expected to answer these questions. So far there have been only indirect observations, apart from detections of some rare objects (e.g., galaxies, quasars and GRBs). The following subsections briefly describe the currently available observational constraints.

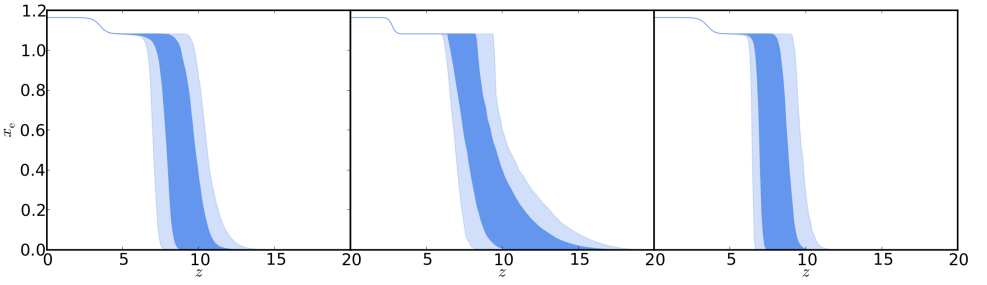

Figure 1.2 – Constraints on the ionization fraction $x_e(z)$ during the EoR from CMB data for three different models. The dark and light blue regions represent the 68% and 95% confidence intervals. Left: $x_e(z)$ as a redshift-symmetric hyperbolic tangent function. Centre: $x_e(z)$ as a redshift-asymmetric power law function. Right: redshift-symmetric parameterization of $x_e(z)$ with additional constraints from the kinetic Sunyaev-Zeldovich effect. Image reproduced from Planck Collaboration et al. (2016a)

1.3.1 Thompson scattering of CMB photons

The primordial CMB radiation is linearly polarized as a result of Thompson scattering between the CMB photons and free electrons before the epoch of recombination. However, the number density of electrons increases due to reionization and causes re-scattering of CMB photons during the EoR. The probability of an interaction between CMB photons and free electrons is related to the Thomson scattering optical depth τ of CMB photons given by:

$$\tau = \int_{t(z)}^{t_0} n_e \sigma_T \, c \, dt, \tag{1.1}$$

where n_e is the number density of free electrons at time t, σ_T is the Thompson scattering cross-section, t_0 is the present time and $t(z)$ is the time at redshift z. The value of n_e depends on the evolution of the neutral hydrogen fraction of the IGM with redshift, or, in other words, on the reionization history. Thompson scattering of the quadrupole component of the CMB radiation field by the free electrons during EoR causes linear-polarization anisotropy in the CMB at large angular scales. The amplitude of the polarization anisotropy is related to the size of the horizon at the redshift when the scattering occurred, and also depends on the scattering optical depth τ. Planck Collaboration et al. (2016a), from the analysis of the CMB temperature and polarization power spectra, found the optical depth for Thompson scattering to be $\tau = 0.058 \pm 0.012$, which can be translated to a model-dependent ionization fraction $x_e(z) = n_e(z)/n_H(z)$, where $n_H(z)$ is the number density of hydrogen atoms. However, this constraint is only sensitive to the line-of-sight integral of $n_e(z)$. Therefore, the reionization history cannot be determined uniquely from the

measurements of the Thompson scattering optical depth alone. Figure 1.2 shows constraints on $x_e(z)$ for three different ionisation models determined using Planck CMB data. Constraints on $x_e(z)$ are currently not very secure and are heavily model dependent.

1.3.2 High redshift quasar and GRB absorption spectra

Absorption spectra of high-redshift quasars can also be used to place constraints on reionization using the Gunn-Peterson effect (Gunn & Peterson 1965). Consider radiation blue-ward of the Ly-α transition, emitted by a background source at a particular redshift. In an expanding universe, the spectrum of the radiation is redshifted and eventually is absorbed by intervening neutral hydrogen clouds along the line of sight at intermediate redshifts. The absorption feature occurs at the frequency corresponding to the rest-frame frequency of Ly-α in the redshifted spectra. After reionization largely completed, the neutral hydrogen density in the IGM is extremely low, and quasar spectra show discrete Ly-α absorption lines due to these (still rare) intervening neutral hydrogen clouds along the sight-line at different redshifts. This collection of absorption lines in a quasar spectrum is called the Ly-α forest.

However, as the IGM becomes more neutral toward higher redshifts (i.e., during the EoR), it absorbs a more significant fraction of the Ly-α emission, developing a 'trough' blue-wards of the Ly-α line in the observed spectra of high redshift quasars. This feature is called the Gunn-Peterson trough and was first observed by Becker et al. (2001) in the high redshift quasar spectra shown in figure 1.3. After this discovery, many optical quasar spectra have been observed to infer the neutral hydrogen fraction in the IGM as a function of redshift (Fan et al. 2003, 2006). A comparison between the optical spectra shown in figure 1.3 suggests that most of the hydrogen in post-reionization Universe ($z < 6$) is ionized and there is a sudden increase in the fraction of neutral hydrogen at $z \gtrsim 6$, suggesting that the EoR ended around $z \approx 6$. However, due to the large scattering cross-section of Ly-α photons with neutral hydrogen, even a minuscule neutral hydrogen fraction ($10^{-4.5}$) can cause the optical depth to saturate (Becker, Bolton & Lidz 2015). Furthermore, the absorption spectra only probe the line of sight evolution of the neutral hydrogen fraction in the IGM and require several observations along different sight-lines to infer its global state. Hence, this method can only be used to place weak lower limits on the IGM neutral fraction.

Similar to quasar absorption spectroscopy, optical spectra of high redshift Gamma-Ray Bursts (GRBs) exhibit Ly-α absorption due to neutral hydrogen in the IGM. The advantage of GRB absorption spectroscopy is that GRBs are very luminous objects and can be detected at even higher redshifts. The first

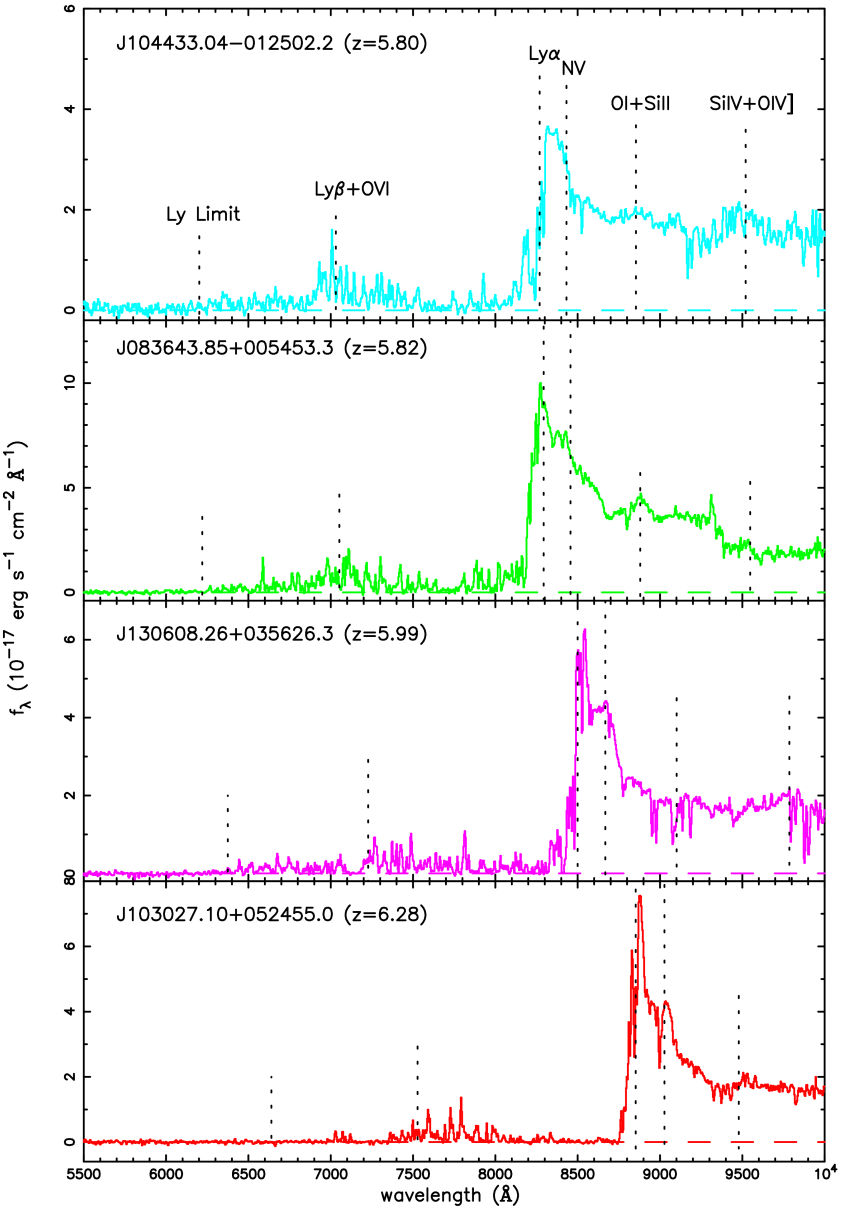

Figure 1.3 – The optical quasar spectra of four high-redshift quasars observed with the Keck telescope. The wavelengths are denoted in the quasar rest frame. The Gunn-Peterson trough becomes more prominent at higher redshifts. Figure reproduced from Becker et al. (2001).

upper limit on neutral hydrogen fraction of $x_{\rm HI} < 0.17$ at $z = 6.3$ was placed by Totani et al. (2006).

1.3.3 High redshift galaxy surveys

Near-infrared observations of high redshift galaxies are instrumental in searching for the sources responsible for reionization. Current and upcoming high-redshift galaxy surveys with the Hubble Space Telescope (Bouwens et al. 2011) and the planned James Webb Space Telescope can and will provide a wealth of information on the nature of these sources. The ultraviolet luminosity functions determined from these surveys can be used to infer the star formation rates at high redshifts. Similarly, observations of GRBs at high redshifts are another tracer of the formation rates of massive stars, and a sufficient number of detections may provide constraints on star formation during the EoR (Wang 2013).

However, inferring the reionization history with galaxy surveys at high redshifts require several assumptions based on models. For example, inferring the star formation rate density from luminosity functions requires assumptions about the initial mass function of stars, which is poorly constrained at high redshifts. Similarly, estimation of the ionization fraction from the star-formation rate density requires information about the escape fraction of Ly-α photons and the IGM recombination rate in the clumpy gas which are again highly uncertain. The lack of knowledge about the parameters mentioned above makes studies of these early phases of the Universe extremely difficult.

1.4 A new probe: the redshifted 21-cm line

Given the limitations of the probes described in section 1.3, only a direct and independent probe of the thermal and ionization history of the IGM during the cosmic dawn reionization can provide a more comprehensive picture of these eras. In recent years, largely thanks to enormous strides forward in low-frequency radio telescope technology and computing, the 21-cm line of neutral hydrogen has become as a promising probe of neutral hydrogen in the IGM during the CD and EoR.

A neutral hydrogen atom, in its ground-state, consists of a single proton and an electron in $1s$ orbital. The electron and proton spins are coupled and can exist in a parallel or anti-parallel spin state, depending on the alignment of the electron and proton spins. These spin-states have an energy difference of about $5.87\,\mu$eV (with the parallel spin state being at higher energy). A (hyperfine) transition between these spin-states can be mediated by absorption or emission of a photon, with a frequency of 1420 MHz (wavelength of about

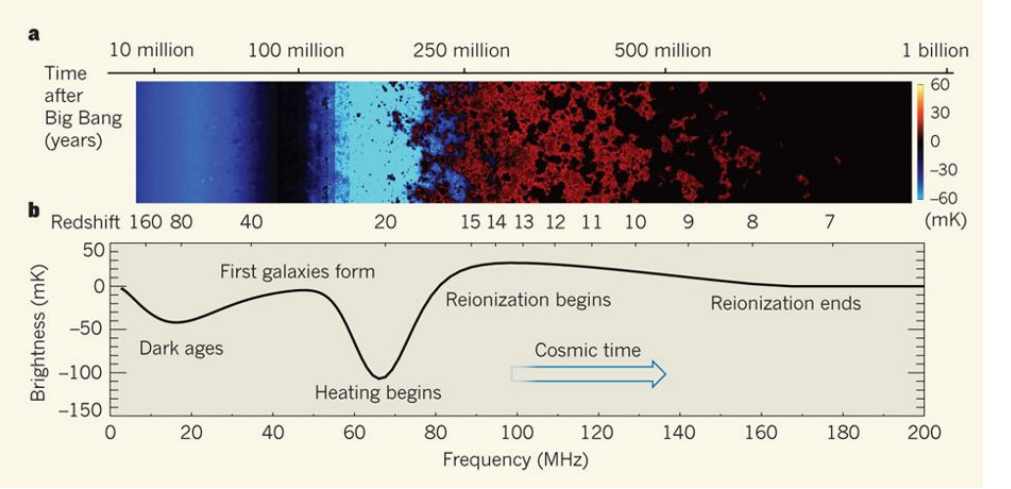

Figure 1.4 – Results of simulations of the expected 21-cm signal as a function of redshift adopted from Pritchard & Loeb (2010a).Top panel: evolution of 21-cm brightness temperature fluctuations during different epochs. Bottom panel: The sky averaged differential brightness as a function of redshift.

21 cm). However, the transition is forbidden, with a transition probability of $2.9\times10^{-15}\,\mathrm{s}^{-1}$ with a lifetime for spontaneous emission of 11 million years. The large abundance of neutral hydrogen (about 75% of all baryonic matter) in the Universe makes the 21-cm transition an excellent tracer of neutral hydrogen in the Universe. Unlike Ly-α photons, 21-cm photons have a smaller chance of getting absorbed by intervening gas in the IGM along the sight-line between the source and the observer. This absorption, in principle, enables three-dimensional tomographic mapping of the Universe during the CD and EoR.

The brightness temperature of the 21-cm line is used to measure the 21-cm signal. The brightness temperature of radiation is equivalent to the temperature of a blackbody in thermal equilibrium with its surroundings. According to Planck's law, the specific intensity I_ν(energy emitted per unit area per unit frequency per unit solid angle) of the radiation emitted by a blackbody, at a temperature T and as a function of frequency, is:

$$I_\nu = \frac{2h\nu^3}{c^2}\frac{1}{e^{\frac{h\nu}{k_BT}}-1}, \tag{1.2}$$

where h is Planck's constant, k_B is Boltzmann's constant, c is the speed of light and ν is frequency. The rest frame frequency of the 21-cm transition is 1420 MHz. The redshifted 21-cm emission from the CD and EoR ($30 \gtrsim z \gtrsim 6$) lies in the wide frequency range of 50-200 MHz. In the limit of very high temperatures or low frequencies (Rayleigh-Jeans approximation), i.e., $h\nu << $

k_BT, Planck's law modifies to:

$$I_\nu = \frac{2\nu^2}{c^2} k_B T = \frac{2k_BT}{\lambda^2} \ . \tag{1.3}$$

The brightness temperature $T_b(\nu)$ can then be expressed in terms of a specific intensity as:

$$T_b(\nu) = \frac{\lambda^2}{2k_B} I_\nu \ . \tag{1.4}$$

The 21-cm transition of neutral hydrogen in the IGM can be observed in absorption or emission against the CMB radiation (Field 1957, 1959). The differential brightness temperature δT_b of the redshifted 21-cm radiation from the CD and EoR depends on several factors such as the brightness temperature of the CMB radiaton (T_{CMB}), the neutral hydrogen fraction x_{HI}, the matter density contrast (δ), and the cosmological parameters. δT_b can be written as (Madau, Meiksin & Rees 1997):

$$\begin{aligned}\delta T_b \approx 27 \ x_{HI} \ (1+\delta) \left(\frac{1+z}{10}\right)^{\frac{1}{2}} \left(1 - \frac{T_{CMB}(z)}{T_S}\right) \left(\frac{\Omega_b}{0.044}\frac{h}{0.7}\right) \\ \times \left(\frac{\Omega_m}{0.27}\right)^{\frac{1}{2}} \left(\frac{1-Y_p}{1-0.248}\right) \left(1 + \frac{1}{H(z)}\frac{dv_\parallel}{dr_\parallel}\right)^{-1} \ \text{mK},\end{aligned} \tag{1.5}$$

where T_S is the spin (or excitation) temperature of the 21-cm transition, Ω_b and Ω_m are the baryonic and total matter density parameters, respectively, Y_p is the primordial helium abundance by mass, h is the normalised Hubble parameter, $H(z)$ is the redshift dependent Hubble parameter and $dv_\parallel/dr_\parallel$ is the proper gradient of the peculiar velocity along the line of sight.

Figure 1.4 shows a simulation of the evolution of the 21-cm brightness temperature fluctuations and sky-averaged brightness temperature with time. The first astrophysical sources appear around $z \sim 30$ and start to ionize the neutral hydrogen in their surroundings. The regions with $\delta T_b = 0$ are the bubbles of ionized gas in the neutral IGM which grow with time and merge to form larger ionized regions. Eventually, the entire IGM gets ionized by $z \sim 6$. The figure also shows the sky-averaged brightness temperature of the 21-cm signal. The evolution profile of the latter shows different time periods when the 21-cm signal is seen in absorption or emission.

1.4.1 Current and planned 21-cm experiments

Several current and next generation interfermoeters, such as the LOw Frequency ARray[1](LOFAR; van Haarlem et al. 2013), the Giant Meterwave Radio

[1] `http://www.lofar.org/`

Telescope[2](GMRT; Paciga et al. 2011), the Murchison Widefield Array[3](MWA; Tingay et al. 2013; Bowman et al. 2013), the Precision Array for Probing the Epoch of Reionization[4](PAPER; Parsons et al. 2010), the Hydrogen Epoch of Reionization Array[5](HERA; DeBoer et al. 2017), NENUFAR[6](New Extension in Nançay Upgrading loFAR; Zarka et al. 2012), and the planned Square Kilometre Array[7](SKA; Mellema et al. 2013; Koopmans et al. 2015) are seeking to statistically detect the power spectrum of the brightness temperature fluctuations in the cosmological 21-cm signal during the EoR and CD.

Complementary to these experiments, single element radiometers such as the Experiment to Detect the Global Epoch of Reionization Signature (EDGES; Bowman et al. 2018), the Large-aperture Experiment to Detect the Dark Ages (LEDA; Bernardi et al. 2016), the Shaped Antenna measurement of the background RAdio Spectrum 2 (SARAS 2; Singh et al. 2017), the Sonda Cosmológica de las Islas para la Detección de Hidrógeno Neutro (SCI-HI; Voytek et al. 2014), the Probing Radio Intensity at high z from Marion (PRIZM; Philip et al. 2018), and the Netherlands-China Low frequency Explorer[8] (NCLE) are aiming to measure the sky-averaged 21-cm signal as a function of redshift.

Although detection of the power spectrum of 21-cm brightness temperature fluctuations remains elusive, several experiments have already placed interesting upper limits on the power spectrum of the 21-cm signal at different redshifts during the EoR. The tightest upper limit has been provided by the LOFAR-EoR Key Science Project (KSP) of $\Delta^2_{21} < (79.6\,\mathrm{mK})^2$ at $k \approx 0.053\,h\,\mathrm{cMpc}^{-1}$ in the redshift range $z = 9.6 - 10.6$ (Patil et al. 2017). Only Ewall-Wice et al. (2016) and Gehlot et al. (2018b) have provided upper limits on the 21-cm signal during the CD using the MWA and LOFAR-Low Band Antenna array, respectively.

Recently, the EDGES collaboration (Bowman et al. 2018) reported a deep spectral feature centred at 78 MHz. This feature was presented as a tentative detection of the long sought-after 21-cm absorption feature seen against the CMB around the redshift $z \sim 17$. Figure 1.5 shows the best-fit absorption profiles obtained with different hardware configurations in the EDGES experiment. The location of this potential absorption trough ($z \sim 17$) is consistent with the simulations and theoretical models of the CD (Pritchard

[2] http://gmrt.ncra.tifr.res.in/
[3] http://www.mwatelescope.org/
[4] http://eor.berkeley.edu/
[5] http://reionization.org/
[6] https://nenufar.obs-nancay.fr/
[7] http://skatelescope.org/
[8] https://www.ru.nl/astrophysics/research/radboud-radio-lab-0/projects/netherlands-china-low-frequency-explorerncle/

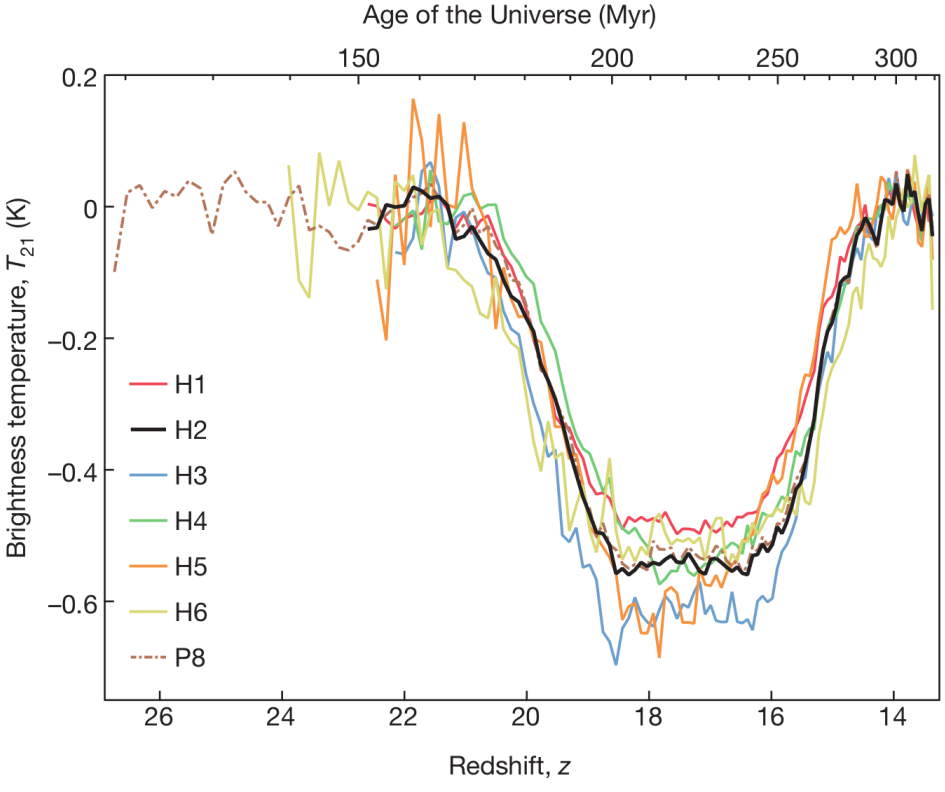

Figure 1.5 – The best-fit 21-cm absorption profiles for different hardware configurations (H1 - H8) of the EDGES experiment in 60-100 MHz frequency range. The black line corresponds to the hardware configuration with the highest signal to noise ratio. The dotted line (P8) corresponds to the same analysis as H2 but with a different foreground model and a wider frequency range. This figure is reproduced from Bowman et al. (2018).

& Loeb 2010a; Mesinger, Ferrara & Spiegel 2013; Cohen et al. 2017). However, the depth ($\Delta T_{21} \sim 0.5\,\mathrm{K}$ within 99% confidence interval) and width ($\Delta\nu \sim 19\,\mathrm{MHz}$) of the feature is considerably deeper and wider than the predictions by the most optimistic models (Pritchard & Loeb 2010b; Fialkov & Barkana 2014; Fialkov & Loeb 2016; Cohen et al. 2017). Although, Hills et al. (2018) raised concerns about this claimed detection, if the detection is verified and confirmed by other ongoing experiments, it might unveil exotic astrophysical phenomena (Barkana 2018; Fialkov, Barkana & Cohen 2018; Ewall-Wice et al. 2018; Feng & Holder 2018; Dowell & Taylor 2018).

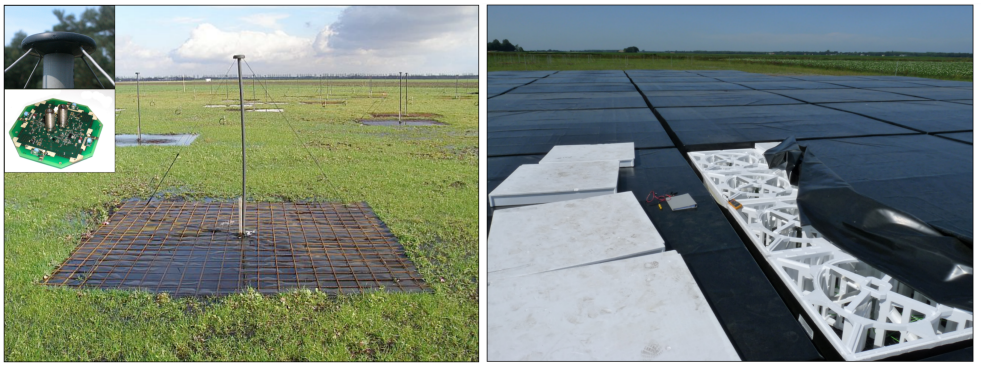

Figure 1.6 – The LBA dipoles (left panel) and the HBA tiles (right panel). Image reproduced from van Haarlem et al. (2013).

1.4.2 The LOw-Frequency ARray (LOFAR)

LOFAR is a state of the art radio interferometer constructed in the northern part of the Netherlands. It operates in the low-frequency range from $10 - 240$ MHz, corresponding to a wavelength range $1.2 - 30$ m, which spans the lowest end of the electromagnetic spectrum observable from Earth. The international LOFAR telescope consists of 51 stations [9], 38 of which are spread across the north-eastern part of the Netherlands and remaining 13 stations are distributed across six countries in Europe, providing baseline lengths ranging from 70 m to 2000 km. Unlike traditional radio telescopes, LOFAR does not have any moving components, the pointing and tracking are achieved by electronically combining the signals from different receiver elements within a station in real-time.

LOFAR design and layout

Each LOFAR station has two distinct types of receivers: Low Band Antenna (LBA) which are designed to operate within $10 - 90$ MHz frequency range, and High Band Antennas (HBA) which are sensitive in $110 - 240$ MHz frequency range. LBAs are low-band dual polarisation droop dipoles having an arm length of 1.38 m placed above a conducting ground plane at an angle of 45° with respect to the ground (left panel of figure 1.6). An LBA dipole has a full-sky Field of View with Full Width Half Maximum (FWHM) of 120°. The HBAs are a group (referred to as a 'tile') of 16 bow-tie shaped dual-polarisation antennas arranged in a 4×4 grid of size 5×5 m (right panel of figure 1.6). The 16 dipoles in an HBA tile are analogue beam-formed to produce a single tile beam with an FWHM of $\sim 25°$ at 150 MHz. The 48 High-Band Antenna

[9] A LOFAR station is akin to a dish in conventional radio telescopes.

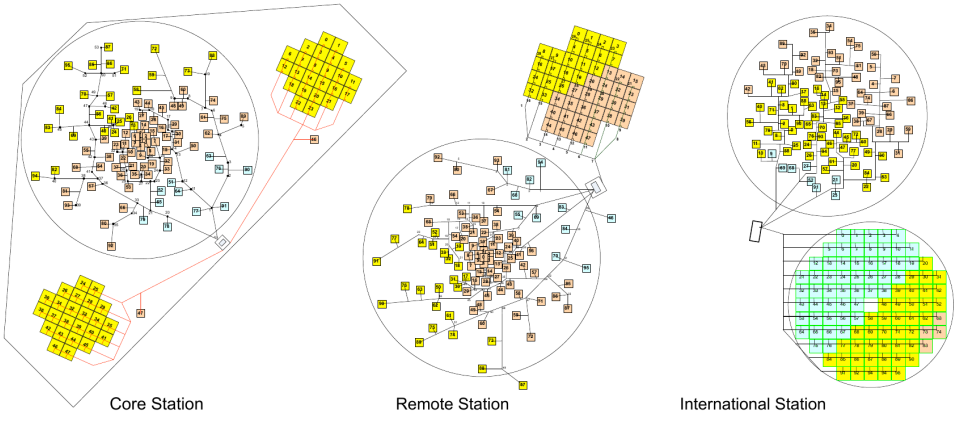

Figure 1.7 – Station layout of Core, Remote and International stations. The big circles represent the LBA dipoles and regular grids of squares represent HBA tiles. Image reproduced from van Haarlem et al. (2013).

tiles of the 24 core stations are split into two independently operating fields.

The LOFAR stations are divided into three categories based on their distance from the centre of the array and receiver layout. Out of 51 LOFAR stations, 24 stations also referred to as the Core Stations, are spread within a dense core of 2 km radius located near Exloo in the Netherlands; 14 stations, also referred to as Remote Stations, are spread across the 40×80 km area in the northeastern part of the Netherlands. The remaining 13 stations called International Stations are distributed across six countries in Europe. The central region of the LOFAR core is called "superterp", and it houses 6 Core Stations within a 300 m diameter. Figure 1.8 shows an aerial photograph of the superterp. The dense core of LOFAR provides unparalleled sensitivity, and the long baselines provide high-resolution imaging at low frequencies.

The main difference between the Core, Remote and International Stations is the number of LBA and HBA dipoles and their layout within the stations. Each Core Station or Remote Station consists of 96 LBA dipoles randomly spread within an area of 81 m diameter and 48 HBA tiles arranged in a regular pattern. However, HBA tile arrangement is different for Core and Remote Stations. The International Stations have 96 LBA dipoles and 96 HBA tiles. The LBA dipole and HBA tile layout for the three stations are shown in figure 1.7.

Figure 1.8 – An aerial of the superterp, located near Exloo (Netherlands). Image reproduced from van Haarlem et al. (2013).

LOFAR signal processing chain

The voltages measured by the receiver elements are digitised using a 12-bit A/D converter with a sampling clock of 200 MHz housed in a receiver unit (RCU). At any given time, the RCUs can only process the signal from 48 out of 96 LBA dipoles or 48 HBA tiles. This setup allows the user to choose from different station configurations, depending on the requirements. The digitised data are then split into 488 sub-bands of 195.3 kHz width by a polyphase filter, providing a total bandwidth of 96 MHz. The digitised signals from the receivers are combined in a beam-former to produce a digitally steerable station beam. The total bandwidth of the system can be traded against the number of beams (pointings) in the sky, e.g., for two pointings, effective bandwidth per pointing is reduced to 48 MHz.

The beam-formed data are transported over a dedicated fibre-optics Wide Area Network (WAN) to the central processing facility (CEP) located at the Center of Information and Technology (CIT; University of Groningen) in Groningen, the Netherlands. The GPU correlator (COBALT) located at the CIT generates correlations (XX, XY, YX, YY) at 1 s integration and 3 kHz frequency resolution (64 channels per sub-band) and stores them in Measurement Set (MS) format on CEP storage clusters. We refer the readers to van Haarlem

et al. (2013) for detailed information of the sensitivity and observing capabilities of the LOFAR international telescope.

1.4.3 The AARTFAAC Cosmic Explorer

If the claimed detection of the 21-cm absorption feature by the EDGES collaboration is validated, the strength of the 21-cm absorption feature will also lead to a substantial increase in the 21-cm brightness temperature fluctuations in the redshift range $z = 16 - 19$ corresponding to the deepest part of the absorption trough (Barkana 2018; Fialkov, Barkana & Cohen 2018). Such large signals make it possible to statistically detect the 21-cm brightness temperature fluctuations in this redshift range on angular scales of about a degree within a ~ 50 times shorter integration time as compared to previous predictions.

This opportunity has motivated us to commence a long-term program called the "AARTFAAC Cosmic Explorer" (ACE) programme, which aims to measure, or set stringent limits on, the power spectrum of the brightness temperature fluctuations of the 21-cm signal from redshift $z \sim 18$ using the Amsterdam-ASTRON Radio Transient Facility And Analysis Centre (AARTFAAC) system (Prasad et al. 2016). AARTFAAC is a wide-field imager based on LOFAR and cross-correlates up to 576 individual receiver elements (LBA dipoles or HBA tiles) from the twelve innermost stations in the LOFAR core, providing a much larger field of view and increased sensitivity on large angular scales compared to the LOFAR beam-formed array. The ACE programme aims for a 1000 h deep integration of a large section of the northern sky to measure or set firm upper limits on the power spectrum of the 21-cm signal. The LOFAR-EoR research group has provisionally been awarded 1000 h of observations with LOFAR-AARTFAAC by the LOFAR time allocation committee to carry out the ACE programme (proposal code: LT10_006). The Principal Investigators of the program are L.V.E. Koopmans and B.K. Gehlot. A brief overview of the LOFAR-AARTFAAC wide-field imager is provided below.

LOFAR-AARTFAAC wide-field imager

AARTFAAC is a LOFAR based all-sky radio transient monitor (Prasad et al. 2016; Kuiack et al. 2018). It piggybacks on currently ongoing LOFAR observations by tapping the digital signal streams from individual receiver elements of a subset of the LOFAR stations and independently processes these signal streams. The choice of the subset of LOFAR stations is defined primarily by sensitivity and image quality. AARTFAAC operates in two modes viz. `A6` where the six innermost stations (the superterp) of the LOFAR core are used, and `A12` where the twelve innermost stations of the LOFAR core are

Figure 1.9 – The layout of AARTFAAC stations. The `A6` mode utilises stations within the white circle. The `A12` mode utilises the `A6` plus six other stations highlighted by white polygons. The white patches represent LBAs and the red patches represent HBAs. This image is adapted from LOFAR interactive map[10].

used. Figure 1.9 shows the stations that are included in AARTFAAC. The `A6` mode consists of 288 receiver elements (LBA dipoles or HBA tiles from six superterp stations) within a 300 m diameter circle and forms a densely sampled uv-plane (almost fully filled uv-coverage within $0-300$ m after a 12-hr track). `A6` is co-planar at the centimetre level, which is optimal for wide-field imaging. The `A12` consists of 576 receiver elements (from 12 central stations of LOFAR core) which provide higher sensitivity and improved resolution imaging over `A6`. The stations utilised in `A12` are chosen such that they provide a uniform uv-coverage while maintaining a reasonably small w-component. For operations in the LBA band, the dipoles in the `LBA_OUTER` station configuration are selected for AARTFAAC observations. In the `LBA_OUTER` configuration, the 48 outermost dipoles are selected from 96 randomly spread dipoles (see van Haarlem et al. (2013) for more information about LOFAR station configurations).

[10] `http://astron.nl/lofartools/lofarmap.html`

The digitised signals from the receiver elements are tapped before digital beamforming and transported to the AARTFAAC GPU correlator located at the CIT in Groningen, the Netherlands. Currently, due to limited network capacity, only 16 sub-bands can be correlated in the 16-bit mode. Each sub-band is 195.3 kHz wide and consists of up to 64 channels, providing a maximum frequency resolution of 3 kHz and a maximum instantaneous bandwidth of 3.1 MHz. The output XX, XY, YX, YY correlations are either dumped on the AARTFAAC storage/compute cluster or routed to the AARTFAAC real-time calibration and the imaging pipeline for transient detection. AARTFAAC can only observe in a drift-scan mode. However, phase-tracking can be applied to the raw data during or after preprocessing. The raw data from the AARTFAAC storage/compute cluster can be streamed via a fast network (1 Gbit/s) to LOFAR processing cluster at the CIT. The raw data can be converted to standard Measurement Set (MS) format using custom software packages. Readers may refer to Prasad et al. (2016) for further information about AARTFAAC system design and capabilities.

1.5 Observational challenges in 21-cm experiments

The 21-cm experiments seeking to measure the faint 21-cm signal from the EoR and CD are limited by real-world challenges, which contaminate and distort this faint signal. The key challenges in 21-cm experiments are contamination due to bright astrophysical foregrounds, distortions of the incoming signals by the ionosphere which is the uppermost ionised part of the Earth's atmosphere, artefacts due to the complex instrumental response, and Radio Frequency Interference (RFI). Overcoming these challenges is the most crucial part of all 21-cm signal experiments, and it is imperative to mitigate or remove these effects in order to detect the faint 21-cm signal.

1.5.1 Foreground emission

The 21-cm signal from the early Universe is extremely faint. According to current theoretical predictions, the brightness temperature of the 21-cm signal ranges from several milli-Kelvin to several tens of milli-Kelvins depending upon the redshift and the spatial scale of the emission. At low frequencies (< 200 MHz), the Galactic foreground emission and extra-galactic sources, such as radio galaxies and clusters, dominate the sky. The intensity of this foreground emission is 4-5 orders of magnitude higher than the expected 21-cm signal. Even the fluctuations on the brightness of these foregrounds are around three orders of magnitude higher than the 21-cm signal.

The foreground emission has two dominant components: (1) diffuse emis-

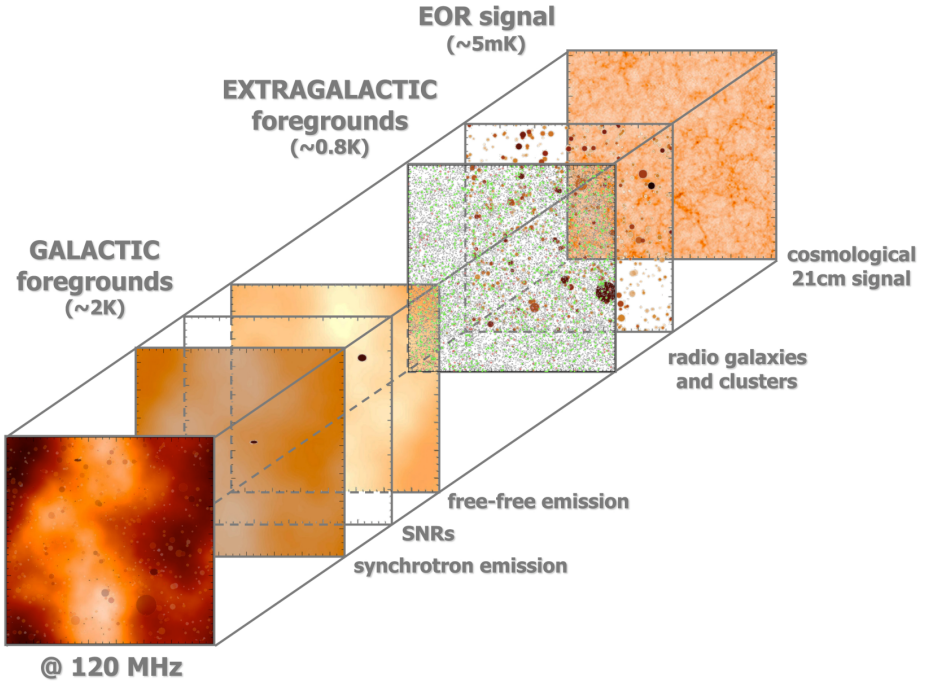

Figure 1.10 – A cartoon depicting the different foreground components that contaminate the 21-cm signal. Image credit: Vibor Jelić

sion comprising Galactic synchrotron emission, Galactic free-free emission and extra-galactic synchrotron emission. The other dominant component is the emission of extra-galactic sources, such as radio galaxies and clusters. Synchrotron emission from supernova remnants (usually within our Galaxy) is another source of foreground emission. Figure 1.10 shows a sketch depicting various foreground components at low frequencies. These foregrounds are mostly a result of continuum emission processes involving the interaction between relativistic charged particles and a magnetic field and have very smooth spectral behaviour which can be modelled using power-laws or higher order log-polynomial models. On the contrary, the 21-cm signal is expected to have a non-smooth spectral structure. All foreground removal techniques exploit these distinct spectral behaviours to remove the foregrounds from the observed data.

1.5.2 Ionospheric effects

Gases in the uppermost part of the Earth's atmosphere are photo-ionised by UV and X-ray radiation from the sun, giving rise to a permanent layer of plasma called the ionosphere. Low-frequency electromagnetic waves propa-

gating through the ionosphere are distorted by fluctuations in the density of the ionospheric plasma. This 'turbulence' induces spatially varying phase-shifts in the propagating wave-front causing refractive and diffractive effects on various spatial scales. These phase-shifts result in refraction and diffraction (e.g., scintillation) of observed sources.

1.5.3 Calibration

Instrumental imperfections, such as complex receiver gains, global band-pass errors, direction-dependent antenna response errors, and ionospheric propagation effects corrupt the sky observed by a radio telescope. These effects must be 'calibrated' in order to achieve a sufficiently high dynamic range while maintaining high image fidelity. Achieving a high level of accuracy in the calibration of the instrument in order to detect the 21-cm signal is one of the most critical challenges in all 21-cm experiments. Self-calibration is the most commonly employed strategy to calibrate the instrument. Self-calibration is an iterative scheme where the difference between the real sky and a 'sky-model' is minimised in order to obtain the instrumental gains and eventually the sky emission. In addition to this, direction-dependent calibration accounts for multiple direction dependent corruption effects during calibration and subtraction of sources.

Nonetheless, there are numerous challenges which limit the accuracy of the calibration desired in 21-cm experiments. For example, the sky-model used during calibration is based on other observations and is often erroneous itself which causes imperfect calibration and subtraction of sources. Furthermore, it is almost impossible to include all sources of foreground emission in the sky-model. Using an incomplete and inaccurate sky-model in calibration leads to biases in the residual data which makes it difficult to obtain the weak 21-cm signal.

Even if the calibration process is perfect and foregrounds are removed with high accuracy, there is the additional thermal-noise uncertainty in the data due to receiver noise and the stochastic nature of the sky emission. This thermal noise is uncorrelated in time and frequency and averages down by integrating over more extended time and larger frequency range. Notwithstanding, several hundred or even thousands of hours of observations are required to reduce the thermal noise and achieve adequate sensitivity to detect the 21-cm signal.

1.6 This thesis

From the previous sections, the reader may have discerned the importance of exploring the Cosmic Dawn and Epoch of Reionization using the 21-cm

signal of neutral hydrogen, to fill the hiatus in our current understanding of the formation and evolution of structure in the Universe. Many ongoing and upcoming efforts aim to probe these eras using low-frequency radio interferometers. These experiments need to overcome significant challenges such as foreground emission, ionospheric distortions, complex instrumental responses, as well as biases generated during the signal processing stage, which all corrupt the weak 21-cm signal.

This thesis presents a range of investigations of essential aspects of these detection challenges and also presents the first 21-cm signal power-spectrum limits at the highest redshifts currently reachable by low-frequency radio telescopes. In particular, the thesis focusses on understanding the consequences of these complex challenges for current and future 21-cm experiments that probe the CD and EoR and developing innovative techniques to overcome, or bypass, these challenges in current and future CD and EoR 21-cm signal experiments.

- Chapter 2 focuses on quantifying (polarised) foregrounds, polarisation leakage, systematic biases and ionospheric effects and their impact on Cosmic Dawn experiments using LOFAR-LBA. Various statistical methods are employed to quantify these effects and discuss methods to mitigate and remove them. The lessons learned from this analysis are valuable in planned experiments with AARTFAAC, NENUFAR, HERA and SKA-Low which all aim to probe the CD.

- Chapter 3 illustrates a novel multi-beam calibration strategy for LOFAR-LBA. It also illustrates the application of the entire LOFAR-EoR processing pipeline on this unique dataset. The lessons learned in chapter 2 help refine the LOFAR-EoR data processing pipeline, which is then used to determine, for the very first time, a 21-cm power spectrum limit in very high redshift range of $z = 19 - 25$, from 14 h of dual-pointing data observed with LOFAR-LBA.

- Chapter 4 provides an overview of the AARTFAAC Cosmic Explorer (ACE) programme, which we have started in order to probe the 21-cm signal brightness temperature fluctuations at redshift $z \sim 18$ corresponding to the location of the alleged 21-cm absorption feature reported by the EDGES collaboration. This chapter also presents the first 21-cm signal power-spectrum results from 6 h of ACE data obtained in a unique "semi drift-scan" mode and processed with the LOFAR-EoR processing pipeline.

- Chapter 5 addresses another critical challenge in 21-cm signal experiments, being the of modelling the diffuse foreground emission and including it in the gain-calibration step. Wide-field observations of a field

around North Celestial Pole with the AARTFAAC-HBA are used to obtain such a model for the first time. The resulting wide-field images are the first-ever images produced with the AARTFAAC-HBA system. Two different methods are demonstrated to model the diffuse emission in the field and compare the results. Finally, an angular power spectrum is presented to quantify the strength of the foreground emission as a function of angular scale.

- Chapter 6 summarises the thesis research and presents a number of concluding remarks, ending with a discussion on future work and follow up research of the presented results.

1